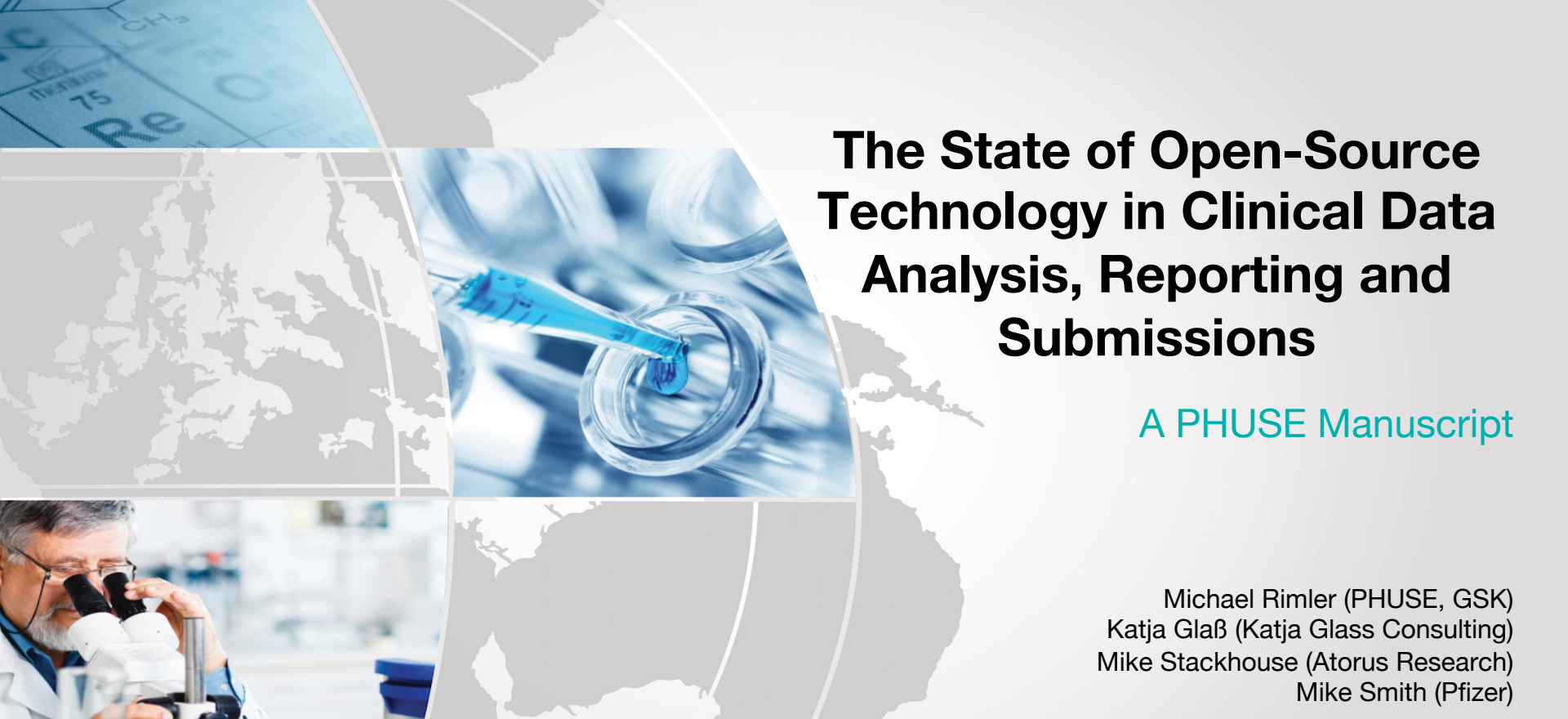




The State of Open-Source Technology in Clinical Data Analysis, Reporting and Submissions

A PHUSE Manuscript

Michael Rimler (PHUSE, GSK)
Katja Glaß (Katja Glass Consulting)
Mike Stackhouse (Atorus Research)
Mike Smith (Pfizer)

2024  phuse us connect



Disclaimer

**These opinions are my own
and do not necessarily reflect
GlaxoSmithKline plc's (GSK)
position or strategy**



The Vision

Purpose

To characterize what we know with respect to ***embedding open-source solutions*** into all matters of clinical data handling, reporting, and submissions



Content

- Broad set of topics
- Comprehensive
 - Quoted content from industry experts
 - Resource lists and links
 - View of the future
- Citable / referenceable answers where possible



Structure

- Introduction
- What have we asked...
 - and answered?
 - and nearly answered?
 - and not yet answered?
- What have we not yet asked?
- Conclusions





Meet the Core Team



Michael Rimler

Open Source Technologies Director, PHUSE



Katja Glaß

*Katja Glass Consulting
COSA Board Member*



Mike Stackhouse

*Chief Innovation Officer, Atorus
PHUSE DVOST Working Group co-Lead*



Mike Smith

Lead of R Centre of Excellence, Pfizer



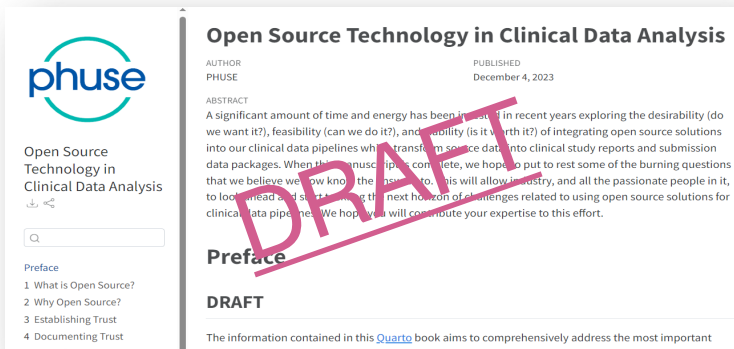
Types of Questions

- Can you trust an open-source solution to be accurate?
- How do you document your trust in an open-source solution to satisfy a third-party inquiry?
- Will the regulatory agencies accept data and analyses generated with solutions developed and available as open source?
- How do you establish reproducibility and traceability with open-source solutions?
- Do we need to match SAS numerically when using a different language?



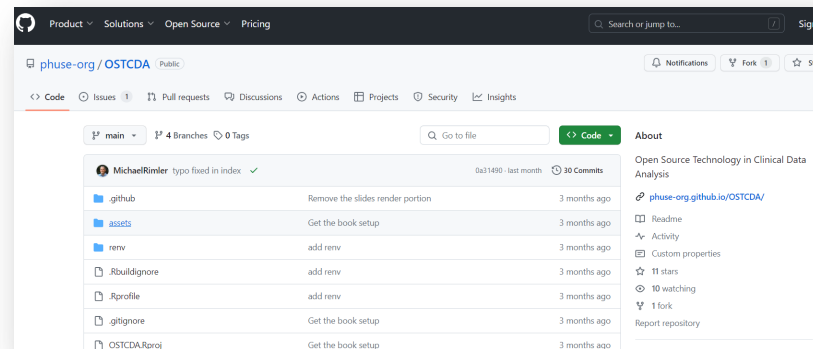
Current Status

Draft manuscript:



<https://phuse-org.github.io/OSTCDA/>

GitHub Repo:



<https://github.com/phuse-org/OSTCDA>

Discussions



How to get involved??



- **Contribute** to the [Discussions](#) in the project repository
- 👍 **Upvote** answers you like
- **Add questions** we've missed
- **Review** the [draft](#) content and submit PR to update content
- Set up a **1:1 interview** michael.rimler@phuse.global

Open Discussion

People

How to...

- ...**transform** the traditional Statistical Analyst in Pharma into the future Data Scientist?

*multilingual
object-oriented
code reviews*

*git/GitHub
software development
agile ways of working*

Tools

Processes

- ...**support users** in managing an ever-evolving environment of open-source packages?

Supporting the People

Transformation



DrLynTaylor

- [PSI Training Course](#) by Andy Nicholls
- [GitHub training video](#) by Christina Fillmore

Discussion
#10

Changing Environment

Discussion
#9

Thank You



**The Global Healthcare
Data Science Community**

Contact Channels

📞 UK +44 1843 609600
✉ office@phuse.global
🌐 phuse.global

Social Media

🐦 [@phusetwitta](https://twitter.com/phusetwitta)
📘 [/phusebook](https://facebook.com/phusebook)
▶ [/phusetube](https://youtube.com/phusetube)
🌐 [/company/phuse](https://linkedin.com/company/phuse)



Backup



Open Discussion

GxP Compliance

Accuracy, Reproducibility, and Traceability

Accuracy

- Can an open-source solution be trusted for accuracy?
- Must an open-source solution match SAS® numerically?

Reproducibility and Traceability

- How do you establish reproducibility and traceability with open-source solutions?
- How do you document your trust in an open-source solution (to satisfy a third-party inquiry)?



GxP Compliance

Accuracy | Reproducibility | Traceability

- R Validation Hub

- [White Paper](#)
- [Validation](#)



- PHUSE E2E OS Guidance

- Comparing Analysis Method Implementations in Software
[\(link to project\)](#)



Packages

- [{testthat}](#)
- [{riskmetric}](#)
- [{logrx}](#)
- [{renv}](#) → renv.lock
- [{valtools}](#)
- [{thevalidatoR}](#)

- PHUSE Blog

*Can Clinical Data Processed with R
be Used in a Regulatory Submission*

Presentations

- R Package Management at Roche ([webinar](#))
- Novo Nordisk Journey to an R based FDA Submission ([webinar](#))
- R Package Validation at Roche ([R/Pharma 2021](#))
- Roche's Approach to Software Validation ([R/Pharma 2022](#))



Suggested docs, appropriately signed-off:

- Overall approach description, including intended use and approach to package dependencies
- Explanation for blanket acceptance of suites of packages
*"We've reviewed {these documents} from R Foundation related to **core R** and accept that they follow good practice."*
- Assessment report for each additional package, likely a combination of human-written interpretation and automated metrics
- Internal processes may govern additional documentation required to qualify/validate the environment.
- Use automated metric tools like {riskmetric} carefully – good information, but do not rely on it too heavily

OSTCDA
Discussion #5